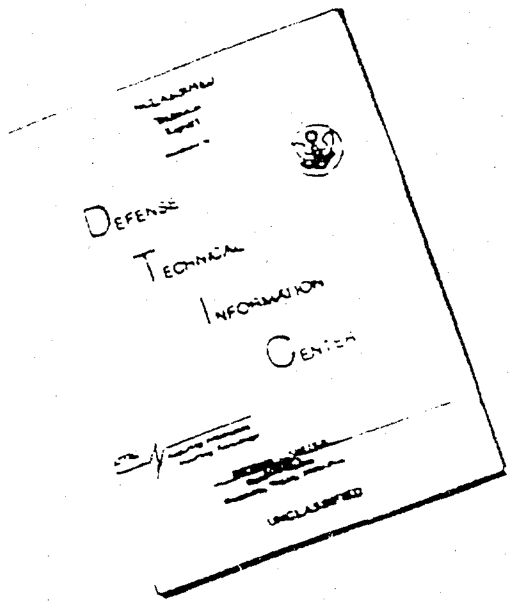
COMPUTING STATIONARY POINTS

B. CURTIS EAVES
MARCH 1977
DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY, CALIFORNIA

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

(6) COMPUTING STATIONARY POINTS.

(9) TECHNICAL REPORT,

(10) B. CURTIS EAVES

(11) Mar [redacted] 77

(12) 32 p.

(15) DAAG29-74-C-0032,
✓ NSF-MPS72-04832

(18) ARO (19) 12391.8-M

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

This research was supported in part by Army Research Office - Durham Contract DAAG-29-74-C-0032 and NSF Grant MPS-72-04832-A03.

402 766
JOB

1. INTRODUCTION

Given a set $\mathcal{X}$ in R^n of form $\{x : Ax \leq a\}$ and a linear function $\mathcal{C}(x) = Cx + c$ from $\mathcal{X}$ to R^n we consider the existence and computation, in a finite number of steps, of a stationary point. A point x^* in $\mathcal{X}$ is defined to be a stationary point of $(\mathcal{C},\mathcal{X})$ if $Cx^* + c$ is an inward normal of $\mathcal{X}$ at x^*, or in other words, if $(y - x^*) \cdot \mathcal{C}(x^*)$ is nonnegative for all choices of y in $\mathcal{X}$.

The existence and computation of stationary points is, in particular, central to the solution of certain quadratic programs, matrix games, and economic equilibrium problems. Such problems can often be cast into the linear complementarity problem which is a stationarity problem $(\mathcal{C},\mathcal{X})$ where $\mathcal{X}$ is the nonnegative orthant; Lemke's algorithm [6] offers the principal avenue for solving the linear complementarity problem.

Herein we adapt Lemke's algorithm in order to approach the general stationarity problem $(\mathcal{C},\mathcal{X})$. Towards describing our main result let x^0 be an arbitrary point in $\mathcal{X}$. We introduce constraints $Bx \leq b$ so that x^0 is the unique solution to $Ax \leq a$ and $Bx \leq b$. Next define $\mathcal{X}_\theta$ to be the set of x's such that $Ax \leq a$ and $Bx \leq b + \theta e$ where $e = (1, 1, \ldots, 1)$ and θ varies over the interval $[0, +\infty)$. Also define a piecewise linear path (X,Θ) to be a (continuous) function $(X,\Theta) : [0, +\infty) \to R^n \times [0, +\infty)$ that is affine on each of a finite number of closed intervals where the closed

intervals cover $[0, +\infty)$. The following theorem captures our principal conclusion.

<u>Theorem</u>: The algorithm computes a piecewise linear path (X,Θ) such that $X(0) = x^0$, $\Theta(0) = 0$, $\Theta(\rho)$ tends to infinity as ρ does, and $X(\rho)$ is a stationary point of $(\mathcal{F}, \mathcal{X}_{\Theta(\rho)})$ for all ρ. ☒

One interesting consequence of the theorem is the following.

<u>Corollary</u>: The algorithm computes either a stationary point x^* of $(\mathcal{F},\mathcal{X})$ or a ray $\{x^* + \theta\bar{x} : \theta \geq \theta^*\}$ in $\mathcal{X}$ such that $\bar{x} \cdot \mathcal{F}(x^* + \theta\bar{x})$ is negative for all θ exceeding θ^*. ☒

For emphasis and clarity we note that throughout the paper C is not assumed to be symetric, $\mathcal{X}$ is not assumed to lie in the nonnegative orthant, and $\mathcal{X}$ is not assumed to have any extreme points.

Some pertinent references are Cottle [1], Lemke [6,7], and [3,4].

2. ALGORITHM

In this section we give a description of the algorithm. Certain details of execution are amplified upon in Section 3, a complete example is executed in Section 4, and the proof of convergence is elaborated upon in Appendices 1 and 2.

Using a theorem of the alternative it is a simple matter to see that finding a stationary point x of $(\mathcal{P},\mathcal{X})$ is equivalent to finding a solution to the system

$$
(1) \qquad \begin{array}{c} Ax + s = a, \qquad Cx + c + A^T\lambda = 0, \\ s \geq 0, \qquad \lambda \geq 0, \qquad s \cdot \lambda = 0 \, . \end{array}
$$

If x is a stationary point, then there is a (s, λ) so that (x, s, λ) is a solution. If (x, s, λ) is a solution then x is a stationary point.

With regard to the data, A is an $m \times n$ matrix, A^T is the transpose of A, a is $m \times 1$, C is $n \times n$, and c is $n \times 1$, and for the variables, x is $n \times 1$, s is $m \times 1$, and λ is $m \times 1$.

Step One of the algorithm is to select any point x^0 in $\mathcal{X}$. $\boxtimes$

In Section 3 we describe a way of executing Step One. If it is noted that $\mathcal{X}$ is a singleton, then the algorithm could be terminated for x^0 is a stationary point, however, it is not necessary to terminate here in this case.

Step Two of the algorithm is to adjoin additional constraints $Bx \leq b$ in order that x^0 is the unique solution to the system.

$$Ax \leq a, \quad Bx \leq b . \qquad \boxtimes$$

Assuming ℓ additional constraints are added, that is, (B, b) is $\ell \times (n + 1)$, then observe that ℓ must be at least as large as, but need not be larger than, $n - h + 1$ where h is the number of linearly independent rows $A_{i.}$ of A such that $A_{i.} x^0 = a_i$. By $(\ )_{i.}$ we denote the ith row of the enclosed matrix. In Section 3 we describe a way of selecting (B, b).

Towards finding a stationary point of $(\mathcal{g}, \mathcal{X})$ we consider the intermediary task of finding stationary points of $(\mathcal{g}, \mathcal{X}_\theta)$ where $\mathcal{X}_\theta$ is the set of x's such that $Ax \leq a$ and $Bx \leq b + \theta e$ where $e = (1, \ldots, 1)$. Clearly $\mathcal{X}_\theta$ is empty for negative θ and bounded for all θ.

Computing a stationary point of $(\mathcal{g}, \mathcal{X}_\theta)$ is equivalent to solving the system,

$$\begin{array}{lll} & Ax + s = a, & Cx + c + A^T\lambda + B^T\mu = 0, \\ (2) & Bx + t - e\theta = b, & s \cdot \lambda = t \cdot \mu = 0, \\ & s \geq 0, \quad t \geq 0, & \lambda \geq 0, \quad \mu \geq 0 . \end{array}$$

Of course, if we have a solution $(x, s, t, \lambda, \mu, \theta)$ to (2) with

$\mu = 0$ then (x, s, λ) solves (1) and x is a stationary point of $(\mathcal{P},\mathcal{Q})$.

Rewriting (2) in detached coefficient form we have

(2)

x	s	t	λ	μ	θ		
C	0	0	A^T	B^T	0		$-c$
A	I	0	0	0	0	$=$	a
B	0	I	0	0	$-e$		b

$$s \geq 0, \quad t \geq 0, \quad \lambda \geq 0, \quad \mu \geq 0, \quad \theta \geq 0,$$

$$s \cdot \lambda = t \cdot \mu = 0 .$$

For purposes of computational efficiency and to emphasis the relation to Lemke's algorithm, we now eliminate x from the system (2). By $(\;)_{\cdot\beta}$ we denote the submatrix of columns indexed by β.

<u>Step Three</u> of the algorithm is to select a set β of n indices j from $\{1, \ldots, m + \ell\}$ so that

$$\begin{pmatrix} A \\ B \end{pmatrix}_{j\cdot} x^0 = \begin{pmatrix} a \\ b \end{pmatrix}_j , \qquad j \in \beta$$

so that the rows

$$\begin{pmatrix} A \\ B \end{pmatrix}_{j\cdot} , \qquad\qquad j \in \beta$$

are linearly independent, and so that

$$\begin{pmatrix} A^T, B^T \end{pmatrix}_{\cdot\beta} \begin{pmatrix} \lambda \\ \mu \end{pmatrix}_{\beta} = -Cx^0 - c$$

$$(\lambda, \mu)_{\beta} \geq 0$$

has a solution. ☒

In Section 3 we describe a way of selecting β.

For any x in $\mathcal{X}_0$ we have

$$x = \begin{pmatrix} A \\ B \end{pmatrix}^{-1}_{\beta\cdot} \left(\begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} I \\ 0 \end{pmatrix} s - \begin{pmatrix} 0 \\ I \end{pmatrix} t + \begin{pmatrix} 0 \\ e \end{pmatrix} \theta \right)_{\beta} . \tag{3}$$

Step Four of the algorithm is to use the expression (3) to eliminate x from the system (2) in order to obtain the system (4) which we have illustrated in detached coefficient form.

(4)

s	t	λ	μ	θ		
J_1	J_2	A^T	B^T	d_1	=	q_1
J_3	J_4	0	0	d_2		q_2

$$s \geq 0,\ t \geq 0,\ \lambda \geq 0,\ \mu \geq 0,\ s \cdot \lambda = t \cdot \mu = 0 .$$

Letting α be those elements in $\{1, \ldots, m + \ell\}$ not in β we have

$$J_1 = -C\begin{pmatrix} A \\ B \end{pmatrix}^{-1}_{\beta\cdot} \begin{pmatrix} I \\ 0 \end{pmatrix}_{\beta\cdot},\quad d_2 = -\begin{pmatrix} 0 \\ e \end{pmatrix}_{\alpha} + \begin{pmatrix} A \\ B \end{pmatrix}_{\alpha\cdot} \begin{pmatrix} A \\ B \end{pmatrix}^{-1}_{\beta\cdot} \begin{pmatrix} 0 \\ e \end{pmatrix}_{\beta},\quad \text{etc.}$$ ☒

Hence (x, s, t, λ, μ) solves (2), if and only if it solves (3) and (4). Thus, if we solve (4) and use (3) to compute x, we have solved (2).

As is done in the simplex method in order to remove coincidence and consequential ambiguity, we perturb the vectors q_1 and q_2 to

$$q_1 + (A^T, B^T)_{\cdot\beta} [\varepsilon]_1 \quad \text{and} \quad q_2 + [\varepsilon]_2$$

where ε is a positive infinitesimal, $[\varepsilon]_1 = (\varepsilon, \varepsilon^2, \ldots, \varepsilon^n)$ and $[\varepsilon]_2 = (\varepsilon^{n+1}, \ldots, \varepsilon^{m+\ell-n})$. The system (4) is thus perturbed to obtain (5).

(5)

s	t	λ	μ	θ	
J_1	J_2	A^T	B^T	e_1	$q_1 + \left(A^T, B^T\right)_{\cdot\beta} [\varepsilon]_1$
J_3	J_4	0	0	e_2	$q_2 + [\varepsilon]_2$

$$s \geq 0,\ t \geq 0,\ \lambda \geq 0,\ \mu \geq 0,\ \Theta \geq 0,\ s \cdot \lambda = t \cdot \mu = 0 \ .$$

That is to say, (3) and (5) are equivalent to (6).

Equivalently, defining $[\varepsilon]_3$ and $[\varepsilon]_4$ by

$$\begin{pmatrix} [\varepsilon]_3 \\ [\varepsilon]_4 \end{pmatrix}_\beta = 0 \qquad \begin{pmatrix} [\varepsilon]_3 \\ [\varepsilon]_4 \end{pmatrix}_\alpha = [\varepsilon]_2$$

(2) is perturbed to (6)

$$
(6)\qquad
\begin{array}{cccccc}
x & s & t & \lambda & \mu & \theta \\
\hline
\multicolumn{1}{|c|}{C} & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{A^T} & \multicolumn{1}{c|}{B^T} & \multicolumn{1}{c|}{0} \\
\hline
\multicolumn{1}{|c|}{A} & \multicolumn{1}{c|}{I} & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{0} \\
\hline
\multicolumn{1}{|c|}{B} & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{I} & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{-e} \\
\hline
\end{array}
\qquad
\begin{array}{|c|}
\hline
-c + \left(A^T,\ B^T\right)_{\cdot\beta} [\varepsilon]_1 \\
\hline
a + [\varepsilon]_3 \\
\hline
b + [\varepsilon]_4 \\
\hline
\end{array}
$$

$$s \geq 0,\ t \geq 0,\ \lambda \geq 0,\ \mu \geq 0,\ \theta \geq 0,\ s \cdot \lambda = t \cdot \mu = 0\ .$$

We observe that $(s_\varepsilon^0,\ t_\varepsilon^0,\ \lambda_\varepsilon^0\ ,\ \mu_\varepsilon^0\ ,\ 0)$ is a solution to (5) where

$$\begin{pmatrix} s_\varepsilon^0 \\ t_\varepsilon^0 \end{pmatrix}_\alpha = \begin{pmatrix} a \\ b \end{pmatrix}_\alpha - \begin{pmatrix} A \\ B \end{pmatrix}_\alpha x_0 + [\varepsilon]_2$$

$$\begin{pmatrix} \lambda_\varepsilon^0 \\ \mu_\varepsilon^0 \end{pmatrix}_\beta = -\left(A^T,\ B^T\right)^{-1}_{\cdot\beta} (Cx^0 + c) + [\varepsilon]_1$$

$$\left(s_\varepsilon^0,\ t_\varepsilon^0\right)_\beta = 0 \qquad \left(\lambda_\varepsilon^0,\ \mu_\varepsilon^0\right)_\alpha = 0$$

This solution of (5) is defined to be the initial solution. We shall show in Appendix 1 that, in fact, the initial solution is the only solution to (5) with $\theta = 0$, note however, that for $\varepsilon = 0$ the solution may not be unique.

The perturbation as described here is a conceptual device; it is implemented in a computer code by making lexicographic comparisons, see Dantzig [2, Chapter 10].

The next step of the algorithm is to "complementary pivot" on (5) à la Lemke wherein one begins with the initial solution and increases θ . To be more explicit here we need some definitions.

Due to the perturbation every solution of (5) has $(m + \ell)$ or $(m + \ell + 1)$ positive components. If a solution has exactly $m + \ell$ positive components it is defined to be a basic solution and the positive variables are called the basic variables. By a ray of solutions to (5) we mean a family of solutions to (5) of form $(s^*, t^*, \lambda^*, \mu^*, \theta^*) + r(\bar{s}, \bar{t}, \bar{\lambda}, \bar{\mu}, \bar{\theta})$ where r varies over the interval $[0, +\infty)$ and $(\bar{s}, \bar{t}, \bar{\lambda}, \bar{\mu}, \bar{\theta})$ is nonzero.

Step Five of the algorithm is to complementary pivot; such is begun in Iteration 1.

Iteration $k = 1$: Beginning with the initial solution, which is basic, increase θ and adjust basic variables in order to retain a solution to (5); either we can increase θ to infinity in this manner or some basic variable, now designated the blocking variable, is driven to zero. In the first case a ray of solutions $\left(s_\varepsilon^0, t_\varepsilon^0, \lambda_\varepsilon^0, \mu_\varepsilon^0, 0\right) + r\left(\bar{s}, \bar{t}, \bar{\lambda}, \bar{\mu}, \bar{\theta}\right)$ to (5) has been generated where the positive components of each solution are θ and the basic

variables; the pivoting is terminated. In the second case a new solution $\left(s_{\varepsilon}^{1}, t_{\varepsilon}^{1}, \lambda_{\varepsilon}^{1}, \mu_{\varepsilon}^{1}, \theta_{\varepsilon}^{1}\right)$ has been generated where the positive components are θ and the basic variables excluding the blocking variable; the new solution is basic and we move to Iteration 2. ☒

Iteration $(k+1) = 2, 3, \ldots$: Assume that the basic solutions $\left(s_{\varepsilon}^{0}, t_{\varepsilon}^{0}, \lambda_{\varepsilon}^{0}, \mu_{\varepsilon}^{0}, \theta_{\varepsilon}^{0}\right), \ldots, \left(s_{\varepsilon}^{k}, t_{\varepsilon}^{k}, \lambda_{\varepsilon}^{k}, \mu_{\varepsilon}^{k}, \theta_{\varepsilon}^{k}\right)$ have been generated and that θ_{ε}^{k} is positive. There is exactly one i in $\{1, \ldots, m+\ell\}$ such that both

$$\begin{pmatrix} s_{\varepsilon}^{k} \\ t_{\varepsilon}^{k} \end{pmatrix}_i = 0 \quad \text{and} \quad \begin{pmatrix} \lambda_{\varepsilon}^{k} \\ \mu_{\varepsilon}^{k} \end{pmatrix}_i = 0 .$$

One of these two variables was basic in the $(k-1)$<u>st</u> basic solution; let us now designate the other variable of the pair, the <u>driving variable</u>. Beginning with the basic solution $\left(s_{\varepsilon}^{k}, t_{\varepsilon}^{k}, \lambda_{\varepsilon}^{k}, \mu_{\varepsilon}^{k}, \theta_{\varepsilon}^{k}\right)$ increase the driving variable and adjust the basic variables to maintain a solution of (5); either we can increase the driving variable to infinity in this manner or some basic variable, now designated the <u>blocking</u> variable, is driven to zero. In the first case a ray of solutions $\left(s_{\varepsilon}^{k}, t_{\varepsilon}^{k}, \lambda_{\varepsilon}^{k}, \mu_{\varepsilon}^{k}, \theta_{\varepsilon}^{k}\right) + r\left(\bar{s}, \bar{t}, \bar{\lambda}, \bar{\mu}, \bar{\theta}\right)$ to (5) has been generated where the positive components of each solution are the driving and basic variables; the pivoting is terminated. In the second case a new solution $\left(s_{\varepsilon}^{k+1}, t_{\varepsilon}^{k+1}, \lambda_{\varepsilon}^{k+1}, \mu_{\varepsilon}^{k+1}, \theta_{\varepsilon}^{k+1}\right)$ has been

generated where the positive components are the driving and basic variables excluding the blocking variable; the new solution is basic and we move to Iteration $k + 2$. ☒

Our description of the algorithm is now essentially complete. The only remaining matter is to show that it generated the path of the Theorem.

We hasten to remark that θ_ε^k is positive for k exceeding zero, because the initial solution is the only one with θ equal to zero and because the algorithm cannot cycle; this arguement is amplified upon in Appendices 1 and 2. Since there are only finitely many basic solutions the complementary pivoting must terminate on a ray.

Assuming the complementary pivoting terminates in iteration $k + 1$, the ray of solutions to (5) generated has the form $\left(s_\varepsilon^k, t_\varepsilon^k, \lambda_\varepsilon^k, \mu_\varepsilon^k, \theta_\varepsilon^k\right) + r(\bar{s}, \bar{t}, \bar{\lambda}, \bar{\mu}, \bar{\theta})$ where $(\bar{s}, \bar{t}, \bar{\lambda}, \bar{\mu}, \bar{\theta}) \neq 0$ and r varies over $[0, +\infty)$.

Define x_ε^k and $\bar{x}$ by

$$x_\varepsilon^k = \begin{pmatrix} A \\ B \end{pmatrix}^{-1}_{\beta\cdot} \left(\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} 0 \\ e \end{pmatrix} \theta_\varepsilon^k - \begin{pmatrix} I \\ 0 \end{pmatrix} s_\varepsilon^k - \begin{pmatrix} 0 \\ I \end{pmatrix} t_\varepsilon^k \right)_\beta$$

$$\bar{x} = \begin{pmatrix} A \\ B \end{pmatrix}^{-1}_{\beta\cdot} \left(\begin{pmatrix} 0 \\ e \end{pmatrix} \bar{\theta} - \begin{pmatrix} I \\ 0 \end{pmatrix} \bar{s} - \begin{pmatrix} 0 \\ I \end{pmatrix} \bar{t} \right)_\beta .$$

Then

$$\left(x_\varepsilon^k, s_\varepsilon^k, t_\varepsilon^k, \lambda_\varepsilon^k, \mu_\varepsilon^k, \theta_\varepsilon^k\right) + r\left(\bar{x}, \bar{s}, \bar{t}, \bar{\lambda}, \bar{\mu}, \bar{\theta}\right)$$

is a ray of solutions to (6) where r varies over the interval $[0, +\infty)$.

Of course, it follows that

(7)
$$\begin{array}{cccccc} \bar{x} & \bar{s} & \bar{t} & \bar{\lambda} & \bar{\mu} & \bar{\theta} \end{array}$$
$$\begin{bmatrix} C & 0 & 0 & A^T & B^T & 0 \\ A & I & 0 & 0 & 0 & 0 \\ B & 0 & I & 0 & 0 & -e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\bar{s} \geq 0 \quad \bar{t} \geq 0 \quad \bar{\lambda} \geq 0 \quad \bar{\mu} \geq 0 \quad \bar{\theta} \geq 0$$

$$\bar{s}\cdot\bar{\lambda} = \bar{t}\cdot\bar{\mu} = \bar{s}\cdot\lambda_\varepsilon^k = \bar{t}\cdot\mu_\varepsilon^k = s_\varepsilon^k\cdot\bar{\lambda} = t_\varepsilon^k\cdot\bar{\mu} = s_\varepsilon^k\cdot\lambda_\varepsilon^k = t_\varepsilon^k\cdot\mu_\varepsilon^k = 0 .$$

Towards proving our main theorem it is important to prove that $\bar{\theta}$ is positive. If the complementary pivoting terminates in iteration one, then obviously $\bar{\theta}$ is positive. If the pivoting terminates in iteration $k + 1$ where k exceeds one, then θ_ε^k is positive as we have already argued. Let us suppose $\bar{\theta}$ is zero and show that it leads to a contradiction.

Suppose $\bar{\theta}$ is zero. Then $\bar{x}$ is zero, since that is the only solution to $Ax \leq 0$ and $Bx \leq 0$. Here $\bar{s}$ and $\bar{t}$ are zero. Therefore, $\bar{\lambda}A + \bar{\mu}B = 0$. Adding $\bar{\lambda}Ax_\varepsilon^k = \bar{\lambda}(a + [\varepsilon]_3)$ to $\bar{\mu}Bx_\varepsilon^k - \bar{\mu}e\theta_\varepsilon^k = \bar{\mu}(b + [\varepsilon]_4)$ we get $-\bar{\mu}e\theta_\varepsilon^k = \bar{\lambda}(a + [\varepsilon]_3) + \bar{\mu}(b + [\varepsilon]_4)$. So if $\bar{\mu}$

is nonzero, the multipliers $(\bar\lambda,\bar\mu)$ show $Ax \leq a + [\varepsilon]_3$ with $Bx \leq b + [\varepsilon]_4$ to be infeasible which is a contradiction. Thus, $\bar\mu$ is zero and $\bar\lambda$ is nonzero. Hence, $\bar\lambda A = 0$ and $\bar\lambda(a + [\varepsilon]_3) = 0$. It follows that $([\varepsilon]_3)_j$ is zero whenever $\bar\lambda_j$ is positive. But this cannot be by choice of $[\varepsilon]_2$, since the rows $A_{j\cdot}$ for $\bar\lambda_j$ positive are linearly dependent. Our supposition leads to a contradiction and we may assume $\bar\theta$ is positive.

Now that $\bar\theta$ is established as positive, we relax our perturbation, that is, for $i = 0, 1, \ldots, k$ the ith solution $(x_\varepsilon^i, s_\varepsilon^i, t_\varepsilon^i, \lambda_\varepsilon^i, \mu_\varepsilon^i, \theta_\varepsilon^i)$ to (6) becomes a solution $(x^i, s^i, t^i, \lambda^i, \mu^i, \theta^i)$ to (2) upon setting ε to zero; a coordinate of the later solution is positive, only if the corresponding solution of the former one is positive. We have that $(x^k, s^k, t^k, \lambda^k, \mu^k, \theta^k) + r(\bar x, \bar s, \bar t, \bar\lambda, \bar\mu, \bar\theta)$ as a ray of solutions to (2).

To prove the Theorem we set $X(i) = x^i$ and $\Theta(i) = \theta^i$ for $i = 0, \ldots, k$, extend (X,Θ) to $[0,k]$ by making it affine on each interval $[i, i+1]$ for $i = 0, \ldots, k-1$, and finally, extend (X,Θ) to $[0, +\infty)$ by setting $X(k+r) = x^k + r\bar x$ and $\Theta(k+r) = \theta^k + r\bar\theta$ for all $r \geq 0$.

To prove the Corollary we first consider μ^k. If μ^k is zero then x^k is a stationary point of $(\mathcal{L},\mathcal{X})$. Otherwise let us assume μ^k is not zero and we consider the ray $x^k + r\bar x$ in $\mathcal{X}$. We have

$$\bar{x} \cdot \mathcal{G}(x^k + r\bar{x}) = -\bar{x} \cdot A^T(\lambda^k + r\bar{\lambda}) - \bar{x} \cdot B^T(\mu^k + r\bar{\mu})$$

$$= \bar{s} \cdot (\lambda^k + r\bar{\lambda}) + (\bar{t} - \bar{\theta}e) \cdot (\mu^k + r\bar{\mu})$$

$$= -\bar{\theta}e \cdot (\mu^k + r\bar{\mu}) < -\bar{\theta}e \cdot \mu^k < 0$$

and the result is established where $x^* = x^k$.

3. SELECTING x^0, (B, b), β

In the previous section the quantities x^0, (B, b) and β were needed to execute the algorithm. In particular settings their selection might be evident, for example, x^0 in $\mathfrak{X}$ might be a known estimate of a stationary point. Our purpose here, however, is to describe a general way of generating x^0, (B, b), and β.

Selecting x^0: Use Phase 1 of the simplex method to solve the system $Ax + Is = a$, $s \geq 0$. Assuming $Ax \leq a$ is feasible, the system (8) is obtained via elementary row operations on $Ax + Is = a$.

$$(8)\qquad \begin{array}{cccc|c} x_\gamma & x_\delta & s_\eta & s_\zeta & \\ \hline I & D_1 & E_1 & 0 & \bar{a}_1 \\ 0 & D_2 & E_2 & I & \bar{a}_2 \geq 0 \end{array}$$

Note that $\bar{a}_2 \geq 0$, that $\{1, \ldots, n\}$ is a disjoint union of γ and δ, that $\{1, \ldots, m\}$ is a disjoint union of η and ζ, and that γ and η contain the same number of elements. Clearly x^0 defined by $x^0_\gamma = \bar{a}_1$ and $x^0_\delta = 0$ is a solution to $Ax \leq a$. ☒

Selecting (B, b): The system

$$(9) \qquad \begin{aligned} A_{\eta\cdot}\ x &= a_\eta \\ x_\delta &\geq 0 \\ -e \cdot A_{\eta\cdot}\ x + ex_\delta &\leq -e \cdot a_\eta \end{aligned}$$

has a unique solution, namely, x^0. To see this observe that (9) is equivalent to

$$A_{\eta\cdot}\ x = a_\eta, \qquad x_\delta = 0$$

and, consequently, to

$$A_{\eta\gamma}\ x_\gamma = a_\eta, \qquad x_\delta = 0\ .$$

But from the elementary row operations we see that $A_{\eta\gamma}^{-1}\ a_\eta = \bar{a}_1$. Therefore, we can define (B, b) by

$$B = \begin{pmatrix} -I_{\delta\cdot} \\ eI_{\delta\cdot} - eA_{\eta\cdot} \end{pmatrix} \qquad b = \begin{pmatrix} 0 \\ -e \cdot a_\eta \end{pmatrix} \qquad \boxtimes$$

<u>Selecting</u> β: Let $\bar{\beta}$ be the set $\eta \cup \{m+1, \ldots, m+\ell\}$ of $n+1$ elements where B is $\ell \times n$. Solve

$$\left(A^T,\ B^T\right)_{\cdot\bar{\beta}} \begin{pmatrix} \lambda \\ \mu \end{pmatrix}_{\bar{\beta}} = -Cx^0 - c$$

$$(\lambda,\ \mu)_{\bar{\beta}} \geq 0$$

for a basic solution. Then discard from $\bar{\beta}$ an element corresponding to a zero component of the solution to get β. $\boxtimes$

4. EXAMPLE

Using the algorithm as specified in Sections 2 and 3 we compute a stationary point of $(\mathcal{L}, \mathcal{X})$ where $\mathcal{L}(x) = Cx + c, \mathcal{X} = \{x : Ax \leq a\}$ and

$$(A, a) = \begin{pmatrix} 1 & 0 & 0 & 2 \\ -1 & 2 & 0 & 6 \\ -2 & -4 & 0 & -4 \end{pmatrix}$$

$$(C, c) = \begin{pmatrix} -1 & 1 & 1 & -2 \\ 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & -1 \end{pmatrix}$$

By projecting $\mathcal{X}$ into the (x_1, x_2) plane we get the set of Figure 1.

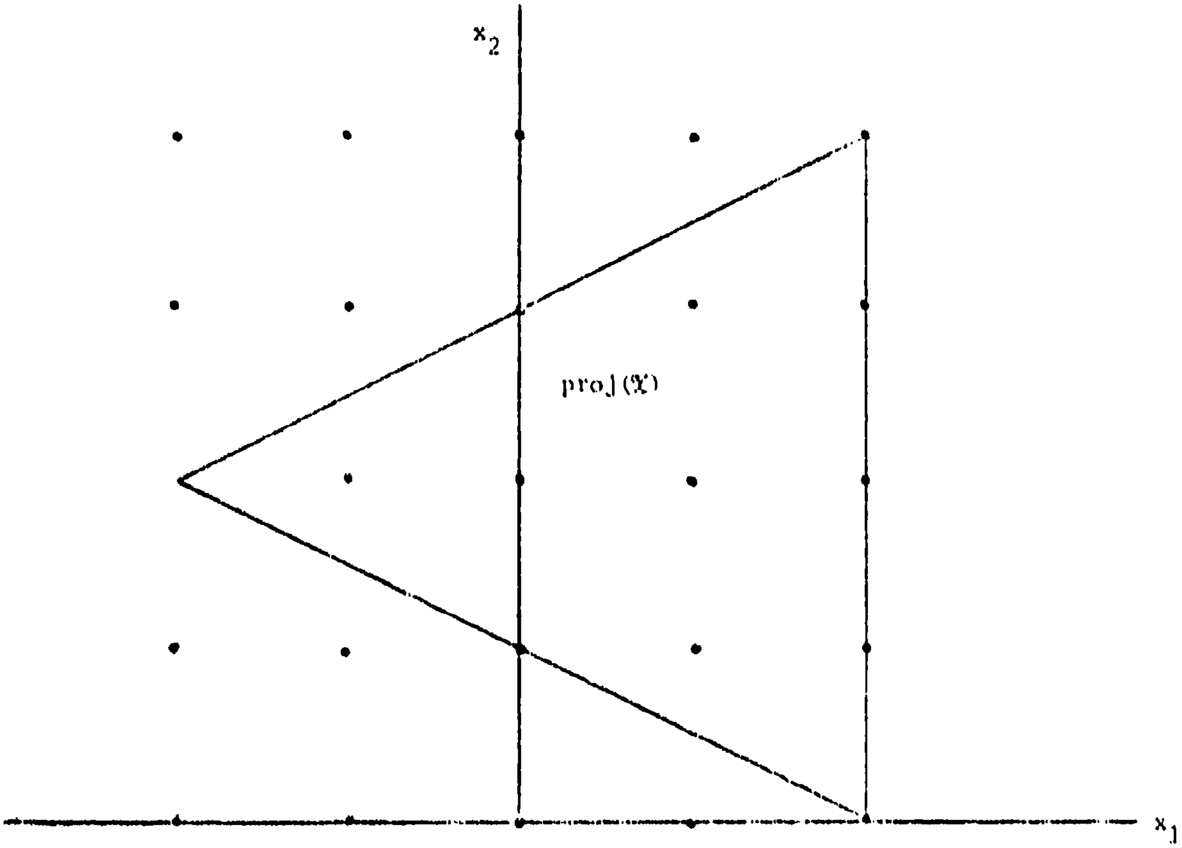


Figure 1

Upon applying Phase I to the system $Ax + Is = a$, $s \geq 0$, that is, the system

x_1	x_2	x_3	s_1	s_2	s_3	
1	0	0	1	0	0	2
-1	2	0	0	1	0	6
-2	-4	0	0	0	1	-4

we get the system (8), namely

x_1	x_2	x_3	s_1	s_2	s_3	
1	0	0	1	0	0	2
-2	0	0	0	1	.5	4
.5	1	0	0	0	-.25	1

whereupon we get $x^0 = (0, 1, 0)$, $\gamma = \{2\}$, $\delta = \{1, 3\}$, $\epsilon = \{1, 2\}$, $\eta = \{3\}$, $\bar{\beta} = \{3, 4, 5, 6\}$, and

$$(B, b) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 3 & 4 & 1 & 4 \end{pmatrix}$$

We calculate $Cx^0 + c = (-1, 0, -1)$ and solve

$$
\begin{array}{cccc}
\lambda_3 & \mu_1 & \mu_2 & \mu_3 \\
\end{array}
$$

$$
\begin{bmatrix} -2 & -1 & 0 & 3 \\ -4 & 0 & 0 & 4 \\ 0 & 0 & -1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}
$$

in nonnegative variables to get $(\lambda_3, \mu_1, \mu_2, \mu_3) = (1, 0, 0, 1)$, and we select $B = \{3, 4, 6\}$. The equation (3) is then

$$
\begin{array}{c|ccccc|}
 & & s_3 & t_1 & t_3 & 0 \\
\hline
x_1 = & 0 & 0 & -1 & 0 & 1 \\
x_2 = & 1 & -.25 & .5 & 0 & -.5 \\
x_3 = & 0 & 1 & 1 & 1 & -2 \\
\hline
\end{array}
$$

and the system (4) is displayed in Figure 2.

	s_1	s_2	s_3	t_1	t_2	t_3	λ_1	λ_2	λ_3	μ_1	μ_2	μ_3	θ	
1	0	0	-.75	-2.5	0	-1	1	-1	-2	-1	0	3	3.5	1
2	0	0	0	1	0	0	0	2	-4	0	0	4	-1	0
3	0	0	-1	-3	0	-1	0	0	0	0	-1	1	4	1
4	1	0	0	1	0	0	0	0	0	0	0	0	-1	2
5	0	1	.5	-2	0	0	0	0	0	0	0	0	2	4
6	0	0	1	1	1	1	0	0	0	0	0	0	-3	0
	1	2	3	4	5	6	7	8	9	10	11	12	13	14

FIGURE 2

Figure 2 displays simultaneously systems (4) and (5). For a given row the perturbation coefficients for $\epsilon, \epsilon^2, \ldots, \epsilon^6$ are found, respectively, in columns 9, 10, 12, 1, 2, 5; this fact remains so following pivots on the matrix.

By pivoting on the positions (1, 9), (2, 10), and (3, 12) of the matrix of Figure 2 the initial solution $(k = 0)$ is displayed, that is, we have the cannonical form with respect to the initial basic solution. Next, for the complementary pivoting we continue pivoting on the matrix on positions (1, 13), (6, 3), (5, 11), (2, 8), (5, 4) and (3, 5) in order to execute iterations $k = 1, 2, \ldots, 6$ whereupon a ray is encountered. The solutions corresponding to each iteration are given in Figure 3; in comumn $\bar{6}$ the values for $(\bar{x}, \bar{s}, \bar{t}, \bar{\lambda}, \bar{\mu}, \bar{0})$ are displayed. As anticipated from the Corollary the algorithm yields a stationary point for $(\not{P}, \mathfrak{X})$, namely, $x = (-1, 2.5, -1)$. In Figure 4 the path $(X,0)$ as projected onto the (x_1, x_2) space is shown.

k	0	1	2	3	4	5	6	$\tilde{6}$
x_1^k	0	-.235	-.8	-1.143	-1.143	-1.1	-1	0
x_2^k	1	1.118	2	2.429	2.429	2.45	2.5	0
x_3^k	0	.471	-.8	-1.143	-1.143	-1.25	-1	0
s_1^k	2	2.235	2.8	3.143	3.143	3.1	3	0
s_2^k	4	3.529	1.2	0	0	0	0	0
s_3^k	0	0	2.4	3.429	3.429	3.6	4	0
t_1^k	0	0	0	0	0	.15	1	1
t_2^k	0	.706	0	0	0	0	1	1
t_3^k	0	0	0	0	0	0	0	1
λ_1^k	0	0	0	0	0	0	0	0
λ_2^k	0	0	0	0	.514	.45	.5	0
λ_3^k	1	0	0	0	0	0	0	0
μ_1^k	0	0	.6	1.286	0	0	0	0
μ_2^k	0	0	0	.429	.172	0	0	0
μ_3^k	1	.059	.2	.286	.029	.05	0	0
θ^k	0	.235	.8	1.143	1.143	1.25	2	1

FIGURE 3

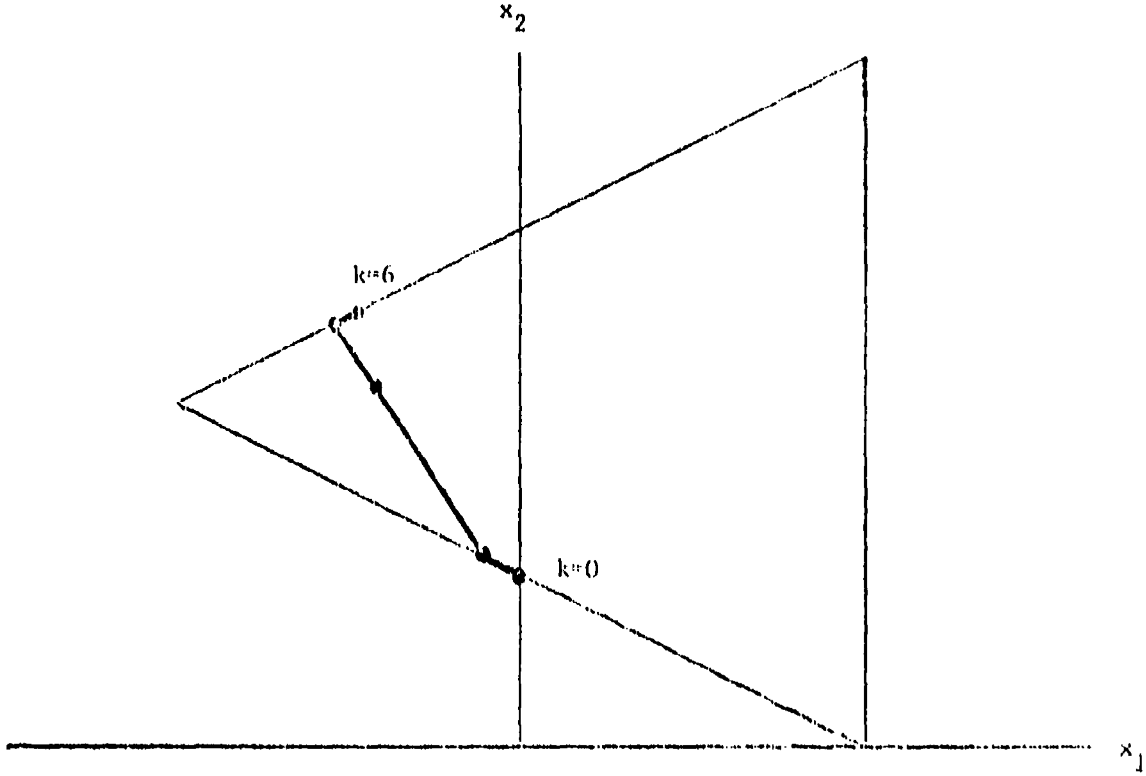


Figure 4

APPENDIX 1: UNIQUE INITIAL SOLUTION

The following lemma shows that (5), or equivalently, (6), has a unique solution with $\theta = 0$ for all sufficiently small positive ε. Observe that it is critical that the constraints are perturbed less than $\wp$. Here our notation has been simplified, and the result is more general than is required in Section 2.

Let $\mathcal{X}_\varepsilon$ be the set of x's in R^n that satisfy $Ax \leq a + B[\varepsilon]_2$ where B is $m \times m$ and $[\varepsilon]_2 = (\varepsilon^{n+1}, \ldots, \varepsilon^{n+m})$. Let $\wp_\varepsilon(x)$ be the linear function $Cx + c + D[\varepsilon]_1$ where D is $n \times n$ and $[\varepsilon]_1 = (\varepsilon^1, \ldots, \varepsilon^n)$. Consider the linear program (y, ε)

$$\min_x \; x \cdot \wp_\varepsilon(y)$$

$$\text{s/t} \quad Ax \leq a + B[\varepsilon]_2$$

Of course, if x is an optimal solution to the program (y, ε) and if $y = x$, then x is a stationary point of $(\wp_\varepsilon, \mathcal{X}_\varepsilon)$.

Lemma: Assume $Ax \leq a$ has a unique solution x^0 and that D is nonsingular. Then there is an X so that $x^0 + X[\varepsilon]_2$ is the unique solution of the linear program (y, ε) for any y in $\mathcal{X}_\varepsilon$ for all sufficiently small positive ε.

Proof: Consider the linear program (x^0, ε). For each (x^0, ε) there is a β such that $A^{-1}_{\cdot\beta}$ exists, $x = A^{-1}_{\beta\cdot}(a + B[\varepsilon]_2)_\beta$ is optimal, and $(A^{-1}_{\beta\cdot})\ \mathcal{P}_\varepsilon(x^0) \leq 0$. Consequently, since there are only finitely many β's and D is nonsingular, there is a β such that $A_\alpha(A^{-1}_{\beta\cdot}(a + B[\varepsilon]_2)_\beta) \leq a_\alpha$ where $\alpha = \sim\beta$ and $(A^{-1}_{\beta})^T\ \mathcal{P}_\varepsilon(x^0) < 0$ for all sufficiently small positive ε. Next observe that ε^{-n} times the diameter of $\mathcal{X}_\varepsilon$ tends to zero as ε tends to zero. Hence, for all sufficiently small positive ε and any y in $\mathcal{X}_\varepsilon$ we have $A_\alpha(A^{-1}_{\beta\cdot}(a + B[\varepsilon]_2)_\beta) \leq a_\alpha$ and $(A^{-1}_{\beta\cdot})^T\ \mathcal{P}_\varepsilon(y) < 0$. It then follows that for all sufficiently small positive ε that $x_\varepsilon = A^{-1}_{\beta\cdot}(a + B[\varepsilon]_2)_\beta$ is the unique optimal solution of the linear program (y, ε) for any y in $\mathcal{X}_\varepsilon$; if $x \neq x_\varepsilon$ is in $\mathcal{X}_\varepsilon$ then $A_{\beta\cdot}x \neq (a + B[\varepsilon]_2)_\beta$, $(\mathcal{P}_\varepsilon(y))^T\ A^{-1}_{\beta\cdot}\ A_{\beta\cdot}x > (\mathcal{P}_\varepsilon(y))^T\ A^{-1}_{\beta\cdot}(a + B[\varepsilon]_2)$, and $x \cdot \mathcal{P}_\varepsilon(y) > x_\varepsilon \cdot \mathcal{P}_\varepsilon(y)$. $\boxtimes$

APPENDIX 2: CONVERGENCE TO A RAY

This appendix is supplied to assist the reader who desires to pursue more carefully the convergence proof for the complementary pivoting of Section 2. Using the concepts of [5] we argue, briefly, that the algorithm generates a ray. A thorough argument is not justified here because it is lengthy and because the literature is now laden with proofs establishing similar conclusions. However, using the remarks made below together with the general theory of [5], one has a rigorous and complete exposition of the convergence to a ray. As in Appendix 1 we simplify the notation.

Consider the system

$$\text{(10)}\qquad \begin{aligned} &Rw + Sz + d\theta = q + Q[\varepsilon] \\ &w \geq 0 \quad z \geq 0 \quad \theta \geq 0 \qquad w \cdot z = 0 \end{aligned}$$

where the matrices R, S, and Q are $n \times n$, Q is nonsingular, the vectors w, z, d, q and $[\varepsilon]$ are $n \times 1$, $[\varepsilon] = (\varepsilon^1, \ldots, \varepsilon^n)$, and the variables are (w, z, θ). Identify this system with (5) where $w = (s, t)$, $z = (\lambda, \mu)$, $q = (q_1, q_2)$, etc.

In Appendix 1 we proved that for each sufficiently small positive ε the system (5) has a unique solution with θ zero, and now let us assume that (10) also has this property.

Let M be the set $R^n \times R^1_+$ and let $\mathcal{M}$ be the subdivision of M where cells of $\mathcal{M}$ are of form $\mathcal{O} \times R^1_+$ and where $\mathcal{O}$ is any orthant of R^n. Define the piecewise linear map $F : M \to R^n$ by

$$F(y,\theta) = \sum_1^n F_i(y_i) + \theta d$$

where $F_i(y_i)$ is defined by

$$F_i(y_i) = \begin{cases} S_{\cdot i} y_i & \text{if } y_i \geq 0 \\ -R_{\cdot i} y_i & \text{if } y_i \leq 0 \end{cases} .$$

First we observe that the system (10) is equivalent to $F(y, 0) = q + [\varepsilon]$, and, in particular, $F(y, 0) = q + [\varepsilon]$ has a unique solution for small positive ε, or in other words, $(F^{-1}(q + [\varepsilon]) \cap (R^n \times 0)$ contains exactly one element where $R^n \times 0$ is the boundary of M.

For small positive ε, $F^{-1}(q + [\varepsilon])$ is a 1-manifold neat in $(M,\mathcal{M})$; let W_ε be the route in $F^{-1}(q + [\varepsilon])$ which meets $R^n \times 0$. The complementary pivot scheme applied to (10) follows W_ε for all sufficiently small ε beginning with the point $W_\varepsilon \cap (R^n \times 0)$. W_ε is subdivided by a finite number of cells of form $W_\varepsilon \cap \sigma$ where σ is an element of $\mathcal{M}$. Since W_ε cannot return to the boundary of M, it must terminate with a ray.

BEST AVAILABLE COPY

BIBLIOGRAPHY

1. COTTLE, R. W., Complementarity and Variational Problems, Instituto Nazionale di Alta Matematica, Symposia Mathematica, Volume XIX (1976), 177-208.

2. DANTZIG, G. B., Linear Programming and Extensions, Princeton University Press, Princeton, N.J. (1963), 621 pp.

3. EAVES, B. C., The Linear Complementarity Problem, Management Science 17 9 (1971), 612-634.

4. EAVES, B. C., On the Basic Theory of Complementarity, Mathematical Programming 1 1 (1971), 68-75.

5. EAVES, B. C., A Short Course in Solving Equations with PL Homotopies, SIAM-AMS Proceedings Vol. 9 (1976), 73-143.

6. LEMKE, C. E., Bimatrix Equilibrium Points and Mathematical Programming, Management Science 11 (1965), 681-689.

7. LEMKE, C. E., Recent Results on Complementarity Problems, Nonlinear Programming, Ed: J. B. Rosen, O. L. Mangasarian, and K. Ritter, Academic Press, New York (1970), 349-384.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ARO 12396.8-M

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 12396.8-M	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Computing Stationary Points		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) B. Curtis Eaves		8. CONTRACT OR GRANT NUMBER(s) DAAG29 74 C 0032
9. PERFORMING ORGANIZATION NAME AND ADDRESS Stanford University Department of Operations Research Stanford University, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE March 1977
		13. NUMBER OF PAGES 29
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Linear equations
Functions (mathematics)
Computation
Stationary points
Economic equilibria
Matrix games
Quadratic programs

≤ or =

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Given a set X in R^n of form $\{x : Ax \leq a\}$ and a linear function $f(x) = Cx + c$ from X to R^n we consider the existence and computation, in a finite number of steps, of a stationary point.

X is C   R to the nth power

DD FORM 1 JAN 73 1473   EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)